Comments on:
ECOFRAM Terrestrial Draft Report
Glenn W. Suter II
U.S. EPA, National Center for Environmental Assessment
June 17, 1999

General Comments:

The Terrestrial ECOFRAM represents a real advance in ecological risk assessment. The authors are to be commended for their good work and OPP is to be commended for taking this step toward genuinely risk-based management of pesticides. As with all reviews, the following addresses primarily points where I believe the document could be improved or clarified. However, my overall positive response should not be forgotten.

The organization and content of the report suggest that the team began by asking what methods could be developed that were probabilistic and then thought about how they might be used (Ch. 6). For example, we learn in Ch. 6 that OPP would still use a tiered assessment approach based on a three-part logic (pass, fail, continue). However, at many points in the text the authors seem to be assuming a two-part logic (pass, fail) or even (pass, continue). An alternative that is also implied at some points is to provide estimates of effects and uncertainties to the risk manager and let him decide what to do. It would be better if the report began with the assessment scheme and then developed the tools to fit the scheme. In particular, the tiers make sense as a way of organizing the methods in a hierarchy of complexity but not necessarily as a way of organizing the assessment process. In general, assessment model complexity should not be the driver because, once you have the new models running, it does not take a lot more time and money to run one set rather than another. However, performing tox tests and other data collection activities does take a lot of time and money. Therefore, it makes more sense to me to organize the tiers in terms of the amount and expense of the data generated. One would then run the models that gave the best estimates of the endpoints given the data that are generated for a particular tier. The assessment may even be more complex when you have less data because of the need to generate estimates. This is not to say that the assessment scheme should be driven by a conventional testing scheme. The testing and measurement should be based on expected improvement in the assessment given current knowledge and expected reductions in uncertainty.

The assessment endpoints are important drivers of the assessment that are treated somewhat superficially in Ch. 2 and then in more detail in Ch. 6, after development of the methods. What expression of endpoints for wildlife does the OPP want? Are endpoints ever probabilities (e.g., the probability of death of an individual organism of a T&E species in the region where the pesticide will be used) or just probabilities of levels of effect? (In fact, the T&E species distinction seems not to have entered into the thought process.) Are endpoints numbers of individuals killed? Are we concerned about individuals if there are mass mortalities but not if they die individually and inconspicuously? When is the endpoint defined as a population response rather than an individual organism response? Are endpoints defined as thresholds for significance

or do you report levels of effect and let the risk manager decide whether it is significant based on the various risk management considerations? These sorts of questions should have been answered *a priori*. Instead it appears that in Ch. 6 you ask what endpoint expressions are generated by the tools that we developed.

The scope of the methods seems to be limited to fully terrestrial birds. Birds that are exposed to pesticides through aquatic food webs are not considered. If they are to be excluded where do you draw the line? For example, you include dermal and drinking water exposure in standing water in fields, but what about foraging in prairie pot holes or wetlands within a pasture or rangeland? Mammals and herps are essentially blown off. While it is true that there are few data for herps, there are a lot of mammalian tox data for any pesticide. What do you do with the rat, mouse, etc. data? If you can extrapolate from *Coturnix* to *Bubo*, you can extrapolate from *Mus* to *Peromyscus*. Also, what about invertebrates, plants, ecosystem processes, etc?

If you use a hypothetical, representative, or other "focal" species, how does that species relate to the endpoint? Is the endpoint really effects on that species or is the species representative of a set of taxonomically, trophically, or behaviorally similar species? If the latter, where is the procedure for doing the translation (e.g., *x* dead robins means 20*x* total dead passerines)? If the former, how do you convey to the risk manager that you effects are much greater than your results suggest?

The concept of comparative risk is not mentioned. A pesticide is not acceptable or unacceptable except in the context of alternatives. Although the output of these methods may not be entirely credible as absolute estimates, they are certainly useful for comparison. However, the expression of results could be different for a comparative assessment than an assessment of absolute acceptability.

There is considerable variability in the extent to which this document provides guidance to the assessor. In the exposure chapter (3), there is a mixture of literature review, recommended methods, justification for methods, and research recommendations. However, few sections contain all of them and they are not presented in a consistent order or format. The effects chapter (4) just presents a lot of methods and leaves it to the following chapter to suggest how they might be used. Hence, the report as a whole is not only short of the goal of sufficient guidance for consistent results, but also is not consistently structured as a guidance document.

There is a certain amount of looking for keys under the lamp post in the treatment of variability and uncertainty. That is, the things that we know well how to estimate, such as residual variance in fitting the dose-response model to the data, get more attention, and some things that are likely to be greater sources of uncertainty get blown off. There is little sense of the relative magnitude of the various sources of variance and uncertainty. For example, given that we are estimating concentrations on food items using Kenaga nomograms, is it really important to consider the difference between a year 2 and year 3 bird? If you were just trying to quantify those sources of variance and uncertainty that can be objectively quantified, you would be OK.

However, if you are serious about calculating the probability that the level of effects will be > *x*, then you can not afford to be less than systematic about your accounting. The same is true if you are really planning to base data demands on the expected reduction in total uncertainty. Accounting for uncertainty is most important when you know the least. Completion of the uncertainty analysis may require use of expert judgement.

The terrestrial ECOFRAM leaves out some routes of exposure (respiratory, dermal, oral due to grooming and preening, etc.) and some modes of action (behavioral and other sublethal effects and indirect effects). While it makes sense to begin with a manageable subset of the problem, if these methods are implemented, the issues that have been left out must somehow be incorporated.

Specific Comments:

p. 2-28 What "valued ecological entity"? The answer seems to be individuals of representative bird species and populations of representative bird species.

Fig. 2.5-1 This is a good generic model, but there seems to be some confusion of exposure **to** a chemical in some medium and exposure **of** a receptor. For example exposure of insects to plant residues leads to both toxicity to insects and residues in insects which leads to exposure to contaminated insects by insectivores; not, exposure to insects leading to residues in insects. The use of terminology leads to confusion about who is exposed to what. It is further confused by the inclusion of exposure processes in some cases and not others. For example, there is toxicity to insects directly from plant residues without an exposure process (i.e., consumption, contact, or both). Also, there is an arrow missing. This is picky, but I believe that in general risk assessors are not sufficiently serious about their conceptual models. It would be useful if you had a true conceptual model that showed the exposures and effects for which you actually have methods.

P. 2-10 The methods may be not easily applied to mammals etc. because of differences in the testing.

p. 2-16 This is a good discussion of the time/effects issues. However, there is some confusion of time to response with duration of exposure. For example, delayed effects of rodenticides seems to be equated with medium term exposure. Also, the issue of appropriate duration is mixed with the issue of the appropriate metric of exposure (dose vs. concentration in food).

P. 3-2 The issue of bioavailability deserves a little more discussion. Bioavailability must be considered when the form of the administered material is significantly different from that to which the organism will be exposed in the field (e.g., dissolved in corn oil vs. incorporated into the tissues of food organisms). Therefore you need relative bioavailability, not absolute, but you need it in all cases where it is an issue.

P. 3-8 It is not clear whether, by "exposure period in the field," you mean time in the field or time in the field engaged in the behavior that leads to exposure.

Sec. 3.3.5 Food consumption seems not to be estimated for critical life stages: nestlings and other young, birds that are refueling during migration, animals that are gorging prior to hibernation, etc.

Sec. 3.3.5 The mixture of foods in the dietary literature may not be relevant to exposures to pesticides, particularly acute exposures. When an animal is in a sprayed field, it consumes the items that are available in that field. During any period of feeding, an animal eats only those items that are available in that season. Animals eat items that are relatively available, such as an abundant pest insect that is falling out of the crop and twitching on the ground. Finally, animals form search images and therefore tend to concentrate on one food type at a time. As a result, it may be realistic to assume that an animal is consuming only one food item during a period of feeding, not just a conservative assumption.

P. 3-31 This chapter talks about conservative assumptions, but assessors really need conservative scenarios to avoid making contradictory assumptions or making an assumption that results in a conservative value of one parameter but not others.

Sec. 3.3.7 It may be important to distinguish aversion from learned avoidance. Also, do altricial nestlings display either one?

Fig. 3.5.1 Some arrows do not connect anything.

Sec. 3.9.3 What is the relevance of skin surface area when an animal is covered with fur or feathers? Is it not likely that oral exposure due to grooming the pelt or preening the feathers will result in much greater exposure than seepage through the fur or feathers to the skin?

Sec. 3.10 What are your criteria for an acceptable transport and fate models? Must they be a mass balance models? The Golder & Associates reference is missing.

P. 4-1, line 11 Why "requires the use of dose-response studies conducted under controlled laboratory conditions using standard laboratory animals?" What about field tests, tests with nonstandard species, monitoring of actual use, etc.?

P. 4-1, lines 26-28 Just because you have tested the species of concern, does not mean that you have estimated the assessment endpoint. The lab-to-field extrapolation, single life-stage to multiple life-stages extrapolation, and the individual to population extrapolations need to be performed.

P. 4-6 The definition of the sublethal factor is unclear.

P. 4-7: After the last bullet, add:

- Have consequent responses resulting in overt effects.
- Have consequent population and community effects.

P. 4-11, lines 1-9: Are long-term dosing studies another option?

P. 4-21, lines 23-26: This makes no sense to me.

Sec. 4.3.2: The authors seem to be saying that they will wait for test development and then see what assessment methods can be developed. It would be better is assessors determined what data they wanted and then demanded that the tests generate those results.

P. 4-34, line 29-30: This is a clear example of the authors saying, we are too uncertain to be uncertain.

P. 4-51, lines13-14: This is another case of uncertainty precluding a probabilistic analysis.

P. 4-56, line 16: One, not zero.

4-65: Why use the 5^{th} percentile of the SSD? How does it relate to the endpoint?

P. 4-75, lines 14-16: Do you think that, because you have only one data point, there is no variance or uncertainty in estimated reproductive effects?

Sec. 4.6: Since the endpoint seems to be individual mortality when using both D-R models and SSDs, the D-R models are actually more sophisticated than the SSDs. The D-R models can estimate the probability of death whereas the SSD only estimates the likelihood of exceeding 50% probability of death?

P. 5-1, lines 14-15: Risk characterization is the component of risk assessment ...

P. 5-8, lines 12-13: What is a multi-species risk assessment model? Are you developing methods for a community endpoint?

P. 5-15: What is the relation of the 10% and 90% values to the endpoints? They seem arbitrary.

Sec. 5.8: Population process models are presented as an alternative to "distribution based models." On the contrary, exposure response distributions are input to the population models. If the OPP plans to do assessments of population effects, they must consider what data are needed and how they will be analyzed.

P. 5-37, line 26-27: What is this overwhelming data limitation? Do we know nothing about avian demographics? Can we not estimate age-specific mortality? Can we not estimate decrements in fecundity? There is uncertainty, but there will always be uncertainty.

Table 6.2-1: One could argue for a long time about what information goes in what tier. For example, I think it is important to get the crop identified early, not leave it until level IV. Don't

we know from the beginning for what crops the pesticide is to be registered? This issue goes back to general comments about using the information that is available, and having a clear rationale for the design of the assessment process.

P. 6-3: Screening need not be computationally simple, just conservative (low Type II error) and use the minimum data set. If you use the quotient method and factors of 10, 100, and 1000 it is as a security blanket while you get accustomed to and efficient with the new methods, not because you need to start that way.